The Medicinal Chemistry Self-Driving Lab - Accelerating Hit Optimisation with a Direct to Biology Approach

Nasim Abdollahi[1], Carla Brown[1], Joseph Brown[1], Stuart Green[1], Lihi Habaz[1], Santha V. Santhakumar[1], Dalia Barsyte[2], Aiymzhan Istayeva[2], Peter Loppnau[2], Magdalena Szewczyk[2], Sarvatit Patel[2], Darshil Patel[3]

1 Acceleration Consortium 2 Structural Genomics Consortium 3 University of Toronto, Department of Chemistry

Correspondence to: Stuart Green stuart.green@utoronto.ca

1. Introduction

The hit-to-lead optimisation phase remains a major bottleneck in the drug discovery pipeline. Although modern screening approaches can rapidly identify weak binders against diverse protein targets, iterative analogue design and synthesis is difficult to scale because purification and downstream testing are time-consuming and hard to automate. The Medicinal Chemistry Self-Driving Lab (MedChem SDL) is being built to address this challenge by developing direct-to-biology workflows that produce actionable biological readouts from compounds present in crude reaction mixtures, enabling faster closed-loop learning during optimisation.

2. Platform and workflow

MedChem SDL is designed to run automated design-make-test-analyze (DMTA) cycles. The design stage uses generative chemistry and activity prediction to prioritise candidate structures. The make stage performs microscale parallel synthesis on a chemistry-capable liquid handler and QC on an analytical LC-MS instrument for confirmation of reaction success. Testing of compounds will occur in a direct to biology manner such as affinity selection mass spectrometry (AS-MS) with the possibility of purification of smaller subset of promising compounds for orthogonal assay confirmation. Finally the analysis stage closes the loop by validating experimental outcomes and improving the underlying models using multi-fidelity data.

A main goal of the MedChem SDL will be to implement AS-MS to rank candidate compounds for protein binding in crude reaction mixtures and using this data to drive the self-driving workflow. In a typical workflow, the target protein is immobilised on beads, incubated to equilibrium with crude reaction mixtures, washed to remove unbound molecules, and bound ligands are eluted and identified/quantified by LC-MS [1-3]. All steps are performed using microtiter plates and plate handling will be performed using a central robotic that interfaces with the synthesis and assay machinery. When complete, all instruments will be orchestrated through a AC-wide opensource orchestration software that will integrate device scheduling and data management.

3. Results and outlook

We assessed the feasibility of direct-to-biology readouts using two cellular assay formats: HiBiT for optimisation of FBXO22 degraders and CETSA for DCAF11 binders. For the CETSA case study analogues were produced around the GW-5074 hit molecule. Compounds demonstrating greater than 120% stabilization in the crude reaction mixtures were purified to benchmark the performance of the crude reaction mixtures against purified material and observed broadly consistent stabilisation trends in crude versus pure compounds at 57 °C. Purified follow-up on a representative hit (compound 125) showed dose-dependent thermal stabilisation and no reduction in cell confluence after 84 h treatment at 10 μM in MCF7, HCT116, and HEK293T cells. We are also have demonstrated the utility of HiBIT assays to optimize the development of degraders based on a FBXO22 recruiter molecule. These results demonstrate the feasibility of using cell-based assays for the optimization of small molecule protein binders.

Ongoing work focuses on integrating cellular assays into the closed loop and applying active learning methods to efficiently navigate chemical space, prioritising experiments that improve both potency/drug-like properties and model learning efficiency.

3.1 Related work

Self-driving laboratories have shown that closed-loop automation coupled with machine learning can accelerate experimental cycles in chemistry and materials research [3–4]. MedChem SDL targets medicinal chemistry hit optimisation, where purification and assay latency are major operational bottlenecks. By combining microscale synthesis with AS-MS and orthogonal cellular assays that can operate on crude mixtures, the platform is intended to increase data density and shorten DMTA iteration time relative to traditional sequential workflows [6].

3.2 Figures and tables

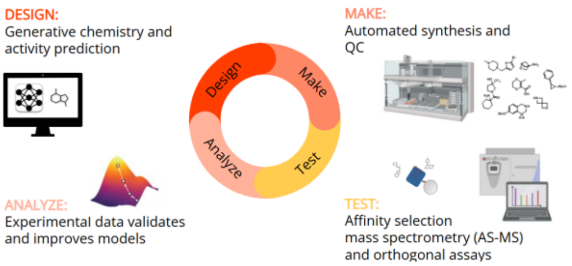

Fig. 1: The Medicinal Chemistry self-driving lab operates in a closed loop experimental workflow to iteratively design, make, test and analyze candidate small molecules based on an initial hit to streamline the early hit-to-lead phase of drug discovery

Fig. 2: All the equipment in the Medicinal Chemistry SDL is accessible by a central robotic arm on a rail and incorporates the necessary equipment for inert atmosphere compound synthesis (Tecan 780 liquid handler), reaction profiling (Waters Acquity LC-MS), bioassay liquid handling (OpenTrons Flex), affinity selection mass spectrometry (Thermo Orbitrap / Vanquish LC-MS), and instruments for orthogonal bioassays (Synergy Neo2 plate reader and QuanStudio Pro 7 thermal Cycler).

Acknowledgments

The Acceleration Consortium is funded by a grant from the Canada First Research Excellence fund (CFREF). Faculty guidance was provided by Prof. Cheryl Arrowsmith (Structural Genomics Consortium) and Prof. Robert Batey (University of Toronto, Department of Chemistry).

References

[1] J. S. Brown et al. Analytical Chemistry (2025), accepted.
[2] A. Quartararo et al. Nature Communications 11 (2020) 3183.
[3] G. Zhang*, J. S. Brown et al. Communications Chemistry 5 (2022).
[4] F. Häse, L. M. Roch, A. Aspuru-Guzik. Next-generation experimentation with self-driving laboratories. Trends Chem. 1(3), 282-291 (2019).
[5] B. Burger et al. A mobile robotic chemist. Nature 583, 237-241 (2020).
[6] R. Prudent, H. Lemoine, J. Walsh, D. Roche. Affinity selection mass spectrometry speeding drug discovery. Drug Discov. Today 28(11), 103760 (2023).